

p#15



RECEIVED

JUN 13 2002

1600

## RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/426,776A

TECH CENTER 1600/2900

DATE: 06/04/2002

TIME: 14:26:20

Input Set : A:\EP.txt

Output Set: N:\CRF3\06042002\I426776A.raw

ENTERED

3 <110> APPLICANT: DING, Jeak Ling  
 4 TAN, Nguan Soon  
 5 HO, Bow  
 6 LAM, Toong Jin  
 8 <120> TITLE OF INVENTION: ISOLATED NUCLEIC ACIDS ENCODING A SECRETORY SIGNAL FOR  
 EXPRESSION AND  
 9 SECRETION OF HETEROLOGOUS RECOMBINANT PROTEINS  
 11 <130> FILE REFERENCE: 1781-0178P  
 13 <140> CURRENT APPLICATION NUMBER: US 09/426,776A  
 14 <141> CURRENT FILING DATE: 1999-10-26  
 16 <160> NUMBER OF SEQ ID NOS: 22  
 18 <170> SOFTWARE: PatentIn version 3.0  
 20 <210> SEQ ID NO: 1  
 21 <211> LENGTH: 29  
 22 <212> TYPE: DNA  
 C--> 23 <213> ORGANISM: Artificial  
 25 <220> FEATURE:  
 26 <223> OTHER INFORMATION: Chloramphenicol acetyltransferase (CAT) gene forward primer  
 derived  
 27 from bacteria  
 29 <400> SEQUENCE: 1  
 30 gaagatctgc tggagaaaaa aatcactgg 29  
 33 <210> SEQ ID NO: 2  
 34 <211> LENGTH: 29  
 35 <212> TYPE: DNA  
 C--> 36 <213> ORGANISM: Artificial  
 38 <220> FEATURE:  
 39 <223> OTHER INFORMATION: Chloramphenicol acetyltransferase (CAT) gene forward primer  
 derived  
 40 from bacteria  
 42 <400> SEQUENCE: 2  
 43 gcatcgcccg tgccttaaaa aaattacgc 29  
 46 <210> SEQ ID NO: 3  
 47 <211> LENGTH: 21  
 48 <212> TYPE: DNA  
 C--> 49 <213> ORGANISM: Artificial  
 51 <220> FEATURE:  
 52 <223> OTHER INFORMATION: OaVtgExon2 reverse primer derived from Oreochromis aureus  
 vitellogenin  
 53 gene exon 2  
 55 <400> SEQUENCE: 3  
 56 ccaagttgga ctggtccccc a 21  
 59 <210> SEQ ID NO: 4  
 60 <211> LENGTH: 19

RECEIVED

JUN 13 2002

TECH CENTER 1600/2900

61 <212> TYPE: DNA  
C--> 62 <213> ORGANISM: Artificial  
64 <220> FEATURE:

## RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/09/426,776A

TIME: 14:26:20

Input Set : A:\EP.txt

Output Set: N:\CRF3\06042002\I426776A.raw

```

65 <223> OTHER INFORMATION: EGFP reverse primer derived from Aequoria victoria green
fluorescent
66     protein
68 <400> SEQUENCE: 4
69 ccctcgccgg acacgctga                                     19
72 <210> SEQ ID NO: 5
73 <211> LENGTH: 29
74 <212> TYPE: DNA
C--> 75 <213> ORGANISM: Artificial
77 <220> FEATURE:
78 <223> OTHER INFORMATION: B-lactamase forward primer derived from bacteria
80 <400> SEQUENCE: 5
81 ccgggatcca gaaacgctgg tgaaagtaa                             29
84 <210> SEQ ID NO: 6
85 <211> LENGTH: 29
86 <212> TYPE: DNA
C--> 87 <213> ORGANISM: Artificial
89 <220> FEATURE:
90 <223> OTHER INFORMATION: B-lactamase reverse primer derived from bacteria
92 <400> SEQUENCE: 6
93 gcggccgtta ccaatgctta atcagtga                             29
96 <210> SEQ ID NO: 7
97 <211> LENGTH: 29
98 <212> TYPE: DNA
C--> 99 <213> ORGANISM: Artificial
101 <220> FEATURE:
102 <223> OTHER INFORMATION: Forward primer from BspSS
104 <400> SEQUENCE: 7
105 gggatcatgag ggtgcttgta ctactctt                             29
108 <210> SEQ ID NO: 8
109 <211> LENGTH: 30
110 <212> TYPE: DNA
C--> 111 <213> ORGANISM: Artificial
113 <220> FEATURE:
114 <223> OTHER INFORMATION: BamGal forward primer with BamHI restriction site and some
beta-
115     galactosidase sequence derived from bacteria
117 <400> SEQUENCE: 8
118 ccatggatcc cgtgatttcg ttgccggtct                             30
121 <210> SEQ ID NO: 9
122 <211> LENGTH: 26
123 <212> TYPE: DNA
C--> 124 <213> ORGANISM: Artificial
126 <220> FEATURE:
127 <223> OTHER INFORMATION: EagGal reverse primer with EagI restriction site
129 <400> SEQUENCE: 9
130 gcgacggccg ggcagacatg gcctgc                                 26
133 <210> SEQ ID NO: 10
134 <211> LENGTH: 21
135 <212> TYPE: PRT
136 <213> ORGANISM: Oreochromis aureus

```

## RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/09/426,776A

TIME: 14:26:20

Input Set : A:\EP.txt

Output Set: N:\CRF3\06042002\I426776A.raw

```

138 <400> SEQUENCE: 10
140 Met Arg Val Leu Val Leu Ala Leu Ala Val Ala Leu Ala Val Gly Asp
141 1          5          10          15
143 Gln Ser Asn Leu Gly
144          20
147 <210> SEQ ID NO: 11
148 <211> LENGTH: 80
149 <212> TYPE: DNA
150 <213> ORGANISM: Oreochromis aureus
152 <220> FEATURE:
153 <221> NAME/KEY: CDS
154 <222> LOCATION: (18)..(80)
155 <223> OTHER INFORMATION: Nucleotides 18-80 encode for SEQ ID NO: 10
157 <400> SEQUENCE: 11
158 attcacatcc accagcc atg agg gtg ctt gta cta gct ctt gct gtg gct          50
159          Met Arg Val Leu Val Leu Ala Leu Ala Val Ala
160          1          5          10
162 ctc gca gtg ggg gac cag tcc aac ttg ggg          80
163 Leu Ala Val Gly Asp Gln Ser Asn Leu Gly
164          15          20
166 <210> SEQ ID NO: 12
167 <211> LENGTH: 204
168 <212> TYPE: DNA
C--> 169 <213> ORGANISM: Artificial
171 <220> FEATURE:
172 <223> OTHER INFORMATION: Junction of Vtgss (derived from Oreochromis aureus) and
CrFCES
173          (Carcinoscorpius rotundicauda ES - EcoRI-SalI flanking fragment of
174          Factor C) determined by sequencing using the Ac5 forward primer and
175          pCDNA3.1/BGH reverse primer
177 <400> SEQUENCE: 12
178 gtggaattct gcagatgcta ccggactcag atcaattcac atccaccagc catgaggggtg          60
179 cttgtactag ctcttgctgt ggctctcgca gtgggggacc agtccaactt gggggatcta          120
180 ggcttggtgt atgaacgag gttcgagtgt aagtgtggcg atccaggcta tgtgttcaac          180
181 attccagtga aacaatgtac atac          204
184 <210> SEQ ID NO: 13
185 <211> LENGTH: 51
186 <212> TYPE: PRT
C--> 187 <213> ORGANISM: Artificial
189 <220> FEATURE:
190 <223> OTHER INFORMATION: VtgCrFCES protein - Vtg derived from Oreochromis aureus and
CrFCES
191          derived from Carcinoscorpius rotundicauda ES - EcoRI-SalI flanking
192          fragment of Factor C
194 <400> SEQUENCE: 13
196 Met Arg Val Leu Val Leu Ala Leu Ala Val Ala Leu Ala Val Gly Asp
197 1          5          10          15
198 Gln Ser Asn Leu Gly Asp Leu Gly Leu Cys Asp Glu Thr Arg Phe Glu
199          20          25          30
200 Cys Lys Cys Gly Asp Pro Gly Tyr Val Phe Asn Ile Pro Val Lys Gln
201          35          40          45

```

## RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/09/426,776A

TIME: 14:26:20

Input Set : A:\EP.txt

Output Set: N:\CRF3\06042002\I426776A.raw

```

202 Cys Tyr Phe
203      50
206 <210> SEQ ID NO: 14
207 <211> LENGTH: 152
208 <212> TYPE: DNA
C--> 209 <213> ORGANISM: Artificial
211 <220> FEATURE:
212 <223> OTHER INFORMATION: Part of the Vtgss-CAT (Vtgss from Oreochromis aureus - CAT
of bacterial
213      origin) fusion in the pBSVtgCAT vector
215 <400> SEQUENCE: 14
216 atcgataagc ttgatgctac cggactcaga tcaattcaca tccaccagcc atgaggggtgc      60
217 ttgtactagc tcttgcgtgtg gctctcgcag tgggggacca gtccaacttg ggggatctgc      120
218 tggagaaaaa aatcaactgga tataccaccg tt      152
221 <210> SEQ ID NO: 15
222 <211> LENGTH: 59
223 <212> TYPE: DNA
C--> 224 <213> ORGANISM: Artificial
226 <220> FEATURE:
227 <223> OTHER INFORMATION: Part of the Vtgss-CAT (Vtgss from Oreochromis aureus - CAT
of bacterial
228      origin) fusion in the pBSVtgCAT vector
230 <400> SEQUENCE: 15
231 ggcggggcgt aattttttta aggcacggcc gatgcgacgg tatcgataac ttgatatcg      59
234 <210> SEQ ID NO: 16
235 <211> LENGTH: 34
236 <212> TYPE: PRT
C--> 237 <213> ORGANISM: Artificial
239 <220> FEATURE:
240 <223> OTHER INFORMATION: Part of the Vtgss-CAT (Vtgss from Oreochromis aureus - CAT
of bacterial
241      origin) fusion in the pBSVtgCAT vector
243 <400> SEQUENCE: 16
245 Met Arg Val Leu Val Leu Ala Leu Ala Val Ala Leu Ala Val Gly Asp
246 1      5      10      15
247 Gln Ser Asn Leu Gly Asp Leu Leu Gln Lys Lys Val Thr Gly Trp Thr
248      20      25      30
249 Thr Val
252 <210> SEQ ID NO: 17
253 <211> LENGTH: 3
254 <212> TYPE: PRT
C--> 255 <213> ORGANISM: Artificial
257 <220> FEATURE:
258 <223> OTHER INFORMATION: Part of the Vtgss-CAT (Vtgss from Oreochromis aureus - CAT
of bacterial
259      origin) fusion in the pBSVtgCAT vector
261 <400> SEQUENCE: 17
263 Gly Gly Ala
264 1
267 <210> SEQ ID NO: 18
268 <211> LENGTH: 66
269 <212> TYPE: DNA
C--> 270 <213> ORGANISM: Artificial

```

## RAW SEQUENCE LISTING

DATE: 06/04/2002

PATENT APPLICATION: US/09/426,776A

TIME: 14:26:20

Input Set : A:\EP.txt

Output Set: N:\CRF3\06042002\I426776A.raw

```

272 <220> FEATURE:
273 <223> OTHER INFORMATION: Part of the nucleotide sequence adjoining Vtgss (derived
from Oreochromis
274 aureus) and CAT (derived from bacteria) in the vector psp-VtgCAT
276 <400> SEQUENCE: 18
277 ggcggggcgt aattttttta aggcacggcc gatgcgacgg tatcgatatt gttacaacac      60
278 cccaac                                          66
281 <210> SEQ ID NO: 19
282 <211> LENGTH: 155
283 <212> TYPE: DNA
C--> 284 <213> ORGANISM: Artificial
286 <220> FEATURE:
287 <223> OTHER INFORMATION: Nucleotide sequence of the Vtg-EGFP (Vtg derived from
Oreochromis
288 aureus - EGFP derived from Aequoria victoria) fusion in the vector
289 pVtgEGFP
291 <400> SEQUENCE: 19
292 gctagcgcta ccggaactcag atcaattcac atccaccagc catgagggtg cttgtactag      60
293 ctcttgctgt ggctctcgca gtgggggacc agtccaactt gggggatcca ccggtcgcca      120
294 ccatggtgag caagggcgtg gtgcagaact ccggg                                          155
297 <210> SEQ ID NO: 20
298 <211> LENGTH: 38
299 <212> TYPE: PRT
C--> 300 <213> ORGANISM: Artificial
302 <220> FEATURE:
303 <223> OTHER INFORMATION: Amino acid sequence of the Vtg-EGFP (Vtg derived from
Oreochromis
304 aureus - EGFP derived from Aequoria victoria) fusion in the vector
305 pVtgEGFP
307 <400> SEQUENCE: 20
309 Met Arg Val Leu Val Leu Ala Leu Ala Val Ala Leu Ala Val Gly Asp
310 1      5      10      15
311 Gln Ser Asn Leu Gly Asp Pro Pro Val Ala Thr Met Val Ser Lys Gly
312      20      25      30
313 Val Val Gln Asn Ser Gly
314      35
317 <210> SEQ ID NO: 21
318 <211> LENGTH: 204
319 <212> TYPE: DNA
C--> 320 <213> ORGANISM: Artificial
322 <220> FEATURE:
323 <223> OTHER INFORMATION: Nucleotide sequence at the junction of Vtgss (derived from
Oreochromis
324 aureus) and B-lactamase (derived from bacteria) in pBADVtgblactKana
326 <400> SEQUENCE: 21
327 ctctactgtt tctccatacc cgtttttttg ggctaacagg aggaattaac catgagggtg      60
328 cttgtactag ctcttgctgt ggctctcgca gtgggggacc agtccaactt gggggatcca      120
329 gaaacgctgg tgaaagtaaa agatgctgaa gatcagttgg gtgcacgagt gggttacatc      180
330 gaactggatc tcaacagcgg taag                                          204
333 <210> SEQ ID NO: 22
334 <211> LENGTH: 51
335 <212> TYPE: PRT
C--> 336 <213> ORGANISM: Artificial

```